

## **Module 10 Lab Activity: Hierarchical Regression as a General Tool for Partitioning Variance**

### **PSY 652 Research Methods**

**Nov 06, 2019**

#### **Description of the dataset:**

*A research team sought to examine factors associated with 21st birthday drinking among female students at a large University. Female students who were nearing age 21 and self-classified as regular drinkers were eligible for the study. In total, 200 students were recruited and agreed to take part in the study. Students were instructed to report to the lab two weeks prior to their 21st birthday. During this lab session, students completed a brief survey that measured alcohol use during the past month (using the Timeline Follow Back Method) and their weight was recorded. One week prior to their 21st birthday, participants were sent a link for an online survey to measure positive alcohol expectancies for drinking on their 21st birthday. Within three days prior to their 21st birthday, students reported to the lab and were given a diary-based data collection form to record several items on their 21st birthday. Students were instructed to record the food that they consumed during the day, the degree to which they were in a partying mood just prior to the celebration, and the quantity and type of drinks that they consumed during the first two hours of the celebration. The students were also given a small breathalyzer machine to measure BAC 2 hours after consumption of their first drink.*

*The dataset called bac\_obs.csv contains the data:*

- *weight: weight in kilograms*
- *alcexp: positive alcohol expectancy for drinking on the impending 21st birthday, a multi-item scale that ranges from 1-7, where a higher score indicates more positive expectations about the role alcohol will play*
- *typ\_drks: the number of standard alcohol drinks consumed in the past 30 days*
- *pmood: a rating on a scale from 1-9 on the respondent's mood to party on the 21st birthday, where 1 means never been less in the mood to party, and 9 means never been more in the mood to party*
- *absorb: a score calculated from the food diaries to determine how full the participant was when they began drinking, the score ranges from 1 to 8, where 1 means a completely full stomach, and 8 means a completely empty stomach*
- *alc\_gm: a score calculated from the drinking diary to estimate the grams of alcohol consumed on the 21st birthday*
- *bac: the participant's blood alcohol content, measured as grams of alcohol per deciliter of blood on the 21st birthday*

1. Download the "bac\_obs.csv" dataset from the Module 10 Lab dropbox folder and save it into a project folder.
2. Create a new R notebook from your project file and name it "bacHR\_notebook"
3. Create a new R chunk with a first level header: "Load Libraries"
  - a. Load the psych, olsrr, and tidyverse packages

4. Create a new R chunk with a first level header: "Import data"
  - a. Read in the "bac\_obs.csv" dataset and assign it to an object named "obs"
5. Create a new R chunk with a first level header: "Describe data"
  - a. Use any method to obtain descriptives for the obs dataset
6. Create a new R chunk with a first level header: "Mutate bac variable"
  - a. In this chunk create a new version of the bac variable called bac100 ( $\text{bac100} = \text{bac} \times 100$ ) and save this to the obs dataframe. We start here because bac is a very small number and multiplying it by a constant won't change the overall model, but will make our output easier to read
  - b. Confirm that this worked by opening the obs dataframe
7. Create a first level header: "Hierarchical regression practice"
8. Create a second level header: "Build Reduced Model A"
  - a. In a new R chunk, regress bac100 on typ\_drks and alc\_gm and obtain model results using the ols\_regress function
9. Create a second level header: "Build Reduced Model B"
  - a. In a new R chunk, regress bac100 on typ\_drks, alc\_gm & weight and obtain model results using the ols\_regress function
10. Create a second level header: "Build Full Model"
  - a. In a new R chunk, regress bac100 on typ\_drks, alc\_gm, weight, & alcexp and obtain model results using the ols\_regress function
11. Create a second level header: "Compare Reduced and Full Models"
  - a. In the white space below, calculate and interpret the unique variance in Y explained by each added predictor. (Hint: think of semi-partial correlation for your interpretations)
  - b. Create a new R chunk with a third level header: "Statistically compare model fit"
    - i. Use the anova function to conduct a partial F-test to statistically compare the three models. List the models in order of increasing size (i.e., most reduced first → full model last).

Hint: `anova(model1, model2, model3, test = "F")`

*The partial F-test tests the null hypothesis that the reduced and full models are not significantly different in terms of model fit. While we can interpret differences in model  $R^2$  simply by comparing these values, the partial F-test allows us to determine if these differences are statistically significant (i.e., if adding a predictor results in a model that explains significantly more variance in Y). The p value for the partial F-test will tell you if there is a significant difference between each model and the model below it.*

- ii. In the white space below, interpret the partial F-test output and answer the following questions.
  1. Does adding the weight variable result in a model that explains significantly more variance in bac100 than Reduced Model A (compare Reduced Model A and Reduced Model B)?

2. Does adding the alc\_exp variable result in a model that explains significantly more variance in bac100 than Reduced Model B (compare Reduced Model B & Full Model)?
3. Reflect on your own research interests and write 2-3 sentences describing an example of when using hierarchical regression could be a good fit for your research. When might you actually use this?